
Bancos de Dados

Aula #1 - Introdução

Prof. Eduardo R. Hruschka

* Slides baseados no material elaborado pelas professoras:

Cristina D. A. Ciferri

Elaine P. M. de Souza

Motivação

- Operações bancárias;
- Compras por intermédio da *Web*;
- Dados sobre produtos/compras em supermercados;
- Dados multimídia;
- Informações geográficas;
- Recuperação de informações na *WWW*;
- Mineração de dados;
- Etc.

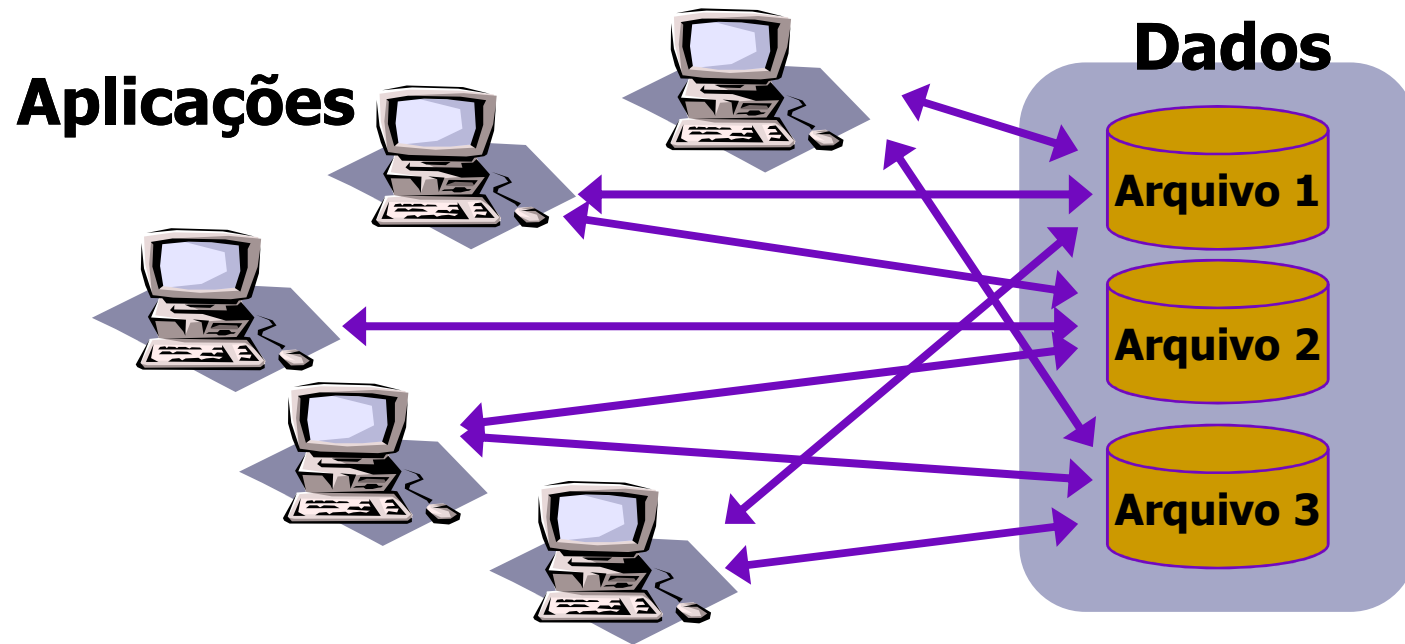
Motivação...

- Necessidade de:
 - armazenar grandes quantidades de dados;
 - acessar informações de maneira eficiente e segura.
- O que é um Banco de Dados (BD) ?
 - Um BD é uma coleção de dados relacionados.¹
 - Dados são fatos que podem ser gravados e que possuem um significado implícito. Por exemplo:
 - Nomes, telefones e endereços de pessoas conhecidas.
 - Uso do termo *banco de dados* é em geral mais restrito, conforme será visto neste curso.

¹ Elmasri & Navathe, Sistemas de Banco de Dados, Pearson, 4ª edição, 2005.

Sistema de Informação (SI) baseado em arquivos:

- ❑ Programas/arquivos orientados a cada unidade organizacional;
- ❑ Rotinas específicas para tarefas específicas;
- ❑ Fortemente dependente da estrutura de dados de cada arquivo.



- ❑ Problemas?

Redundância:

**Aplicação de
Produção**



**Arquivos de Dados
de Produção**



**Aplicação de
Vendas**



**Arquivos de Dados
de Vendas**



**Aplicação de
Compras**



**Arquivos de Dados
de Compras**



Inconsistência:

Aplicação de Produção



Inserir:

Nome: Notebook
NroSerie:1111111
Fabricante: Y



Aplicação de Vendas



Inserir:

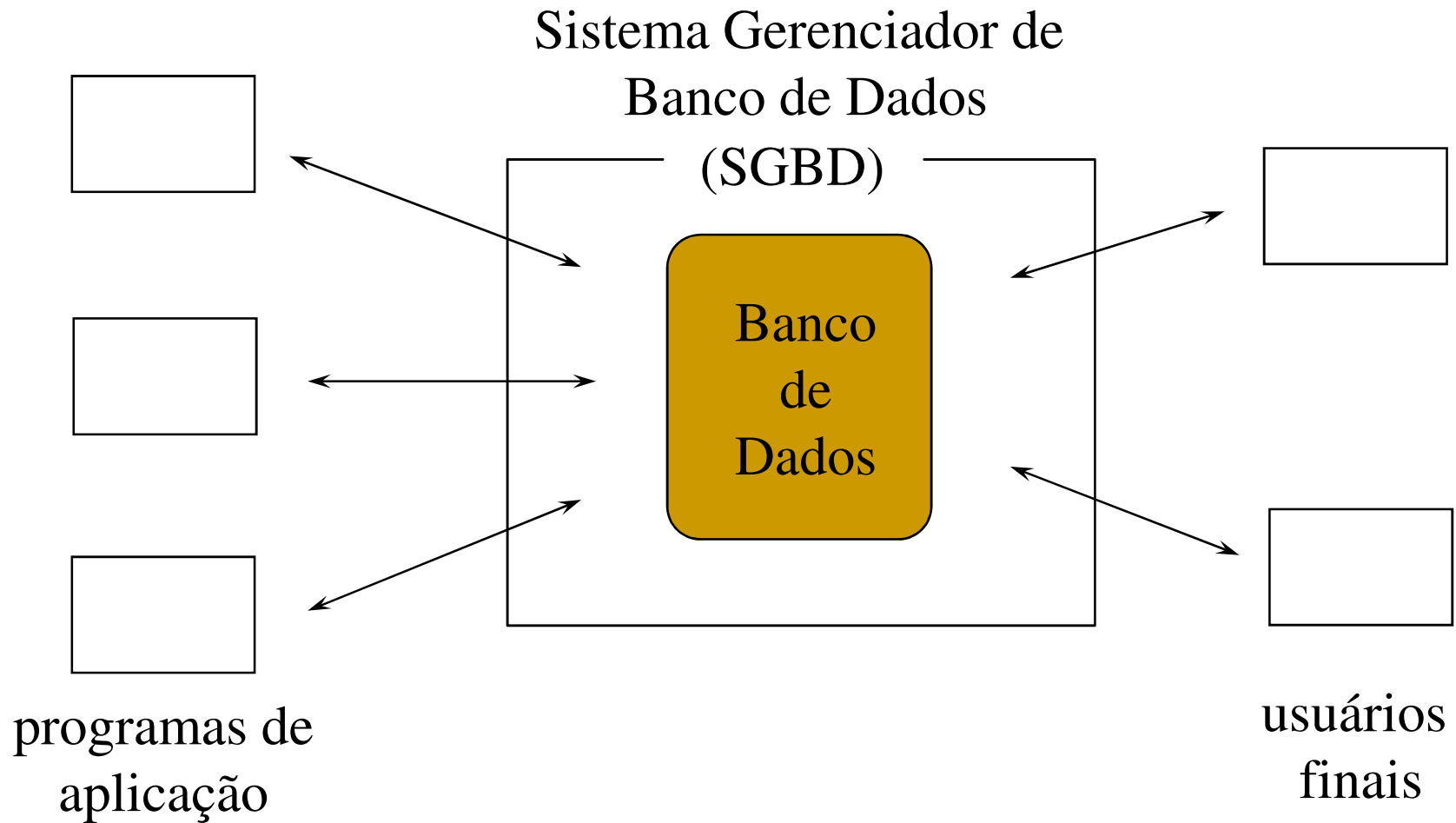
Nome: Notebook
NroSerie:1111111
Fabricante: X



Outros problemas de SI baseado em arquivos:

- Dificuldade de acesso aos dados;
 - Isolamento dos dados;
 - Falta de garantias de integridade;
 - Dificuldades para garantir atomicidade;
 - Anomalias no acesso concorrente;
 - Comprometimento da segurança;
 - Excessivamente dependente das estruturas de dados comuns aos programas:
 - alterações nestas estruturas → modificações nos programas.
-

Sistema de Banco de Dados (SBD)

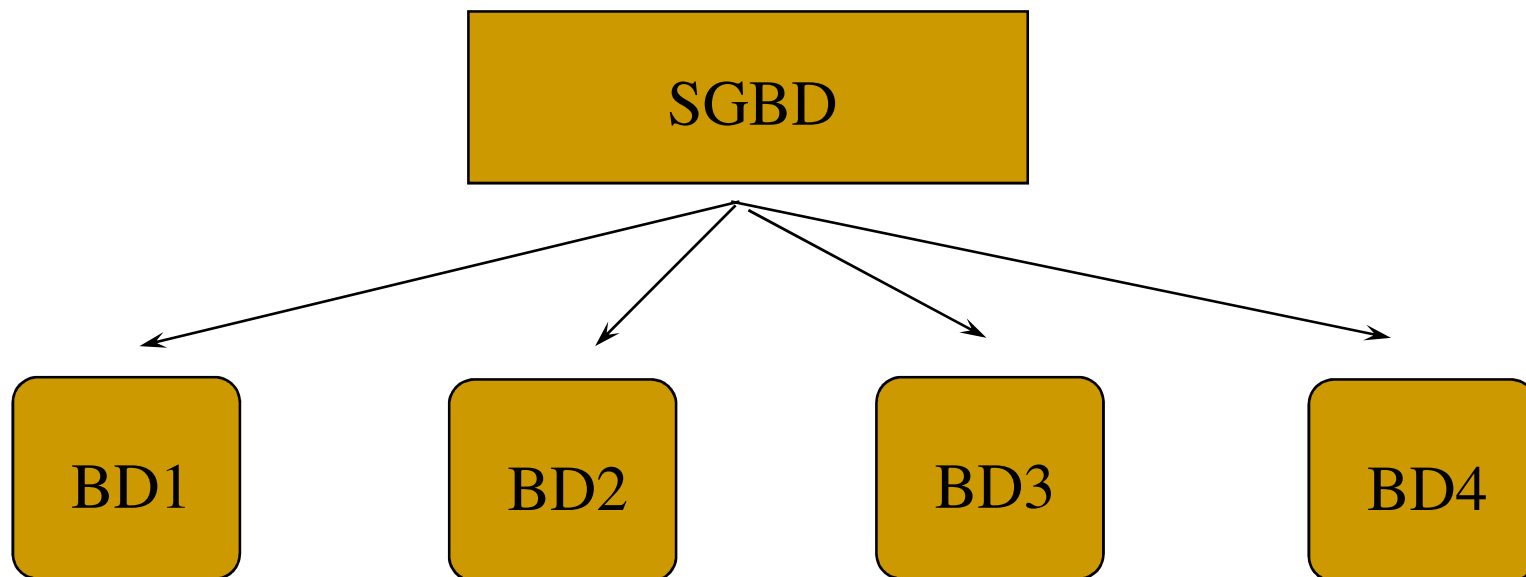


Sistema de Banco de Dados (SBD)

- Sistema de armazenamento de dados:
 - ❑ Manter informações;
 - ❑ Torná-las disponíveis quando necessário;
 - ❑ Armazenamento não volátil.
- Componentes:
 - ❑ Banco de dados (BD);
 - ❑ Sistema Gerenciador de Banco de Dados (SGBD);
 - ❑ Usuários;
 - ❑ Hardware;

Banco de Dados (BD)

- Depósito de dados armazenados;
- Os dados devem ser logicamente coerentes;
- Uma coleção randômica não é um BD.



Sistema Gerenciador de Banco de Dados (SGBD):

- *Conjunto* de dados associado a um conjunto de programas para acesso a estes dados:
 - Criar o BD;
 - Manter o BD.
- Propósito de armazenar, de maneira eficiente, grandes quantidades de informações, permitindo ao usuário buscar e atualizar tais informações;
- Camada existente entre os dados e os usuários;
- Isola os usuários dos detalhes de *hardware*;

Recursos de um SGBD

- adição de novos arquivos;
- inserção de dados;
- recuperação de dados;
- atualização dos dados;
- eliminação dos dados;
- criação de visões;
- atribuição de privilégios, etc.;

Usuários

- Administrador do BD (*Database Administrator – DBA*):
 - coordena e monitora o uso do BD (segurança, tempo de resposta ruim, etc.);
 - autoriza o acesso ao BD;
 - *ajustar* projeto físico quando necessário;
 - adquire *software* e *hardware* necessários;
- Projetista do BD:
 - identificar os dados a serem armazenados no BD;
 - escolher as estruturas apropriadas para representar e armazenar esses dados;
 - levantar necessidades dos usuários (requisitos).

Usuários

- Programador de aplicações:
 - escreve os programas aplicativos;
 - realiza requisições ao SGBD;
- Usuário final:
 - manipula o BD através de:
 - linguagens de consulta;
 - programas previamente desenvolvidos;
 - tipos de usuários:
 - Leigos (operador de terminal) X sofisticados (engenheiros);
 - Casuais (gerentes) X freqüentes (caixas de bancos);

Hardware

- Volumes de armazenamento secundário;
- Dispositivos de entrada e saída;
- Controladores de dispositivos;
- Processador, memórias associadas;
- Etc.

Vantagens da Utilização de SGBD

- Redundância controlada:
 - Redundância:
 - mesmos dados armazenados várias vezes:
 - Abordagem tradicional baseada em processamento de arquivos de uma universidade;
 - Dados sobre alunos na secretaria acadêmica e na contabilidade:
 - Duplicação de esforços ao *cadastrar* dados;
 - Espaço de armazenamento;
 - Inconsistências (inserções, eliminações, atualizações, etc.)
 - Consistência dos dados armazenados
 - Inconsistência:
 - quando dados duplicados armazenam valores distintos;
 - existe quando a redundância não é controlada;
 - redundância controlada pode permitir recuperações mais eficientes.
-

Vantagens da Utilização de SGBD

- **Segurança:**
 - com relação ao acesso ao sistema:
 - *login* dos usuários;
 - com relação ao acesso aos dados do sistema:
 - visões parciais, de acordo com os usuários;
 - acesso controlado, através de graus de privilégios:
 - Informações confidenciais, consultas, atualizações, etc.
 - **Prover funcionalidades para a definição e a garantia das restrições de integridade:**
 - garantem a precisão dos dados (tipos de dados *inteiro*, *real*, *caractere*, ...);
 - relacionamentos entre registros (disciplina vinculada a curso);
 - especificam as restrições impostas pelo sistema real: *salário=0*, ...
-

Vantagens da Utilização de SGBD

- Compartilhamento de dados por vários usuários:
 - base de dados é definida apenas uma vez.
 - Padronização:
 - formato dos dados (tipos);
 - domínio dos valores dos dados.
 - Existência de diferentes interfaces:
 - linha de comando, gráfica.
 - Representação de relacionamentos entre os dados;
 - Recuperação de falhas de *software* e *hardware*;
 - Facilidade de desenvolvimento de aplicações: SGBD *retira* dos programas aplicativos a complexidade de gerenciamento de estruturas de acesso aos dados.
-

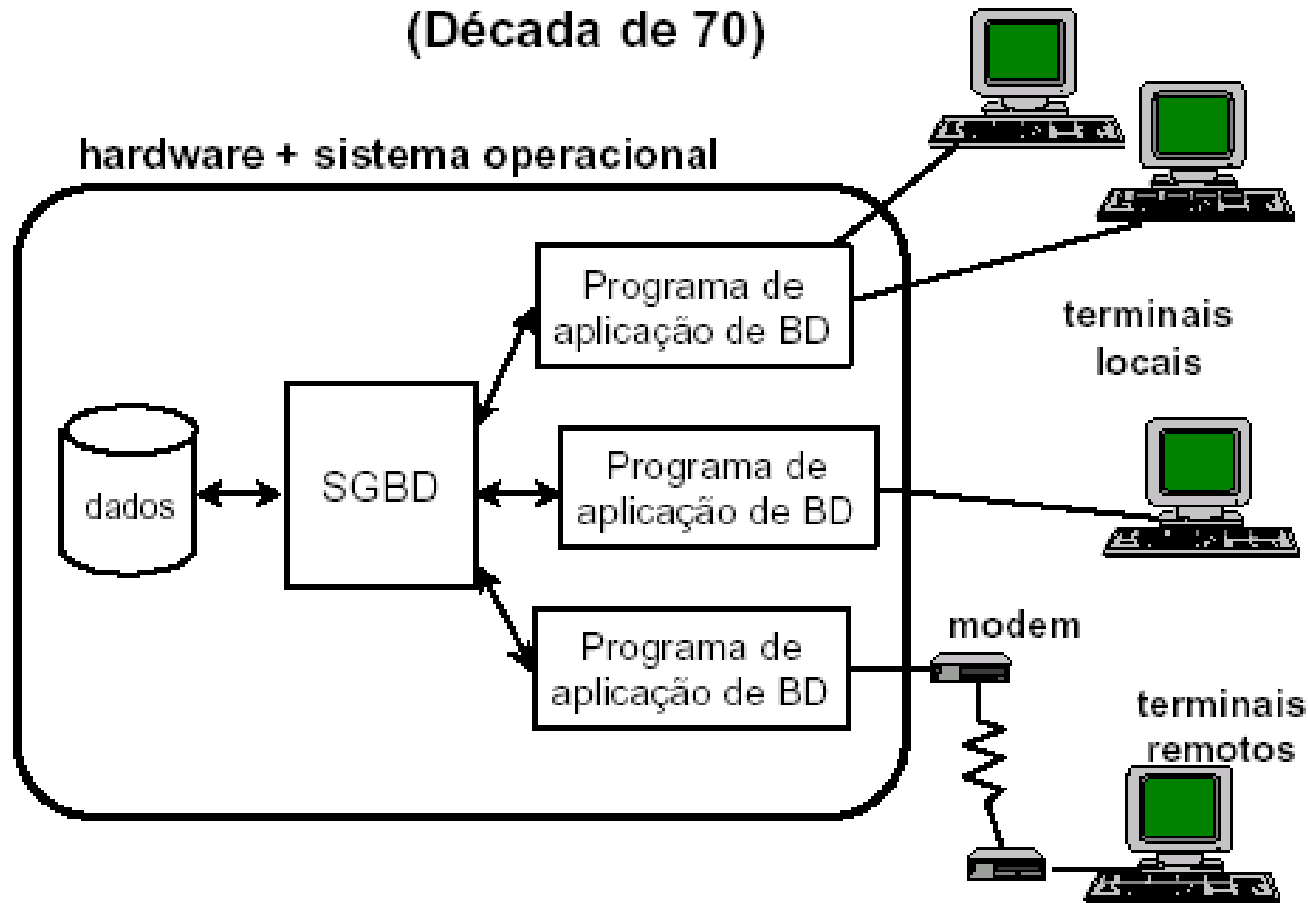
Desvantagens de um SGBD

- A generalidade para definição e processamento de dados (segurança, controle de concorrência, recuperação, funções de integridade, etc.) origina altos custos:
 - ❑ Software;
 - ❑ Hardware;
 - ❑ Treinamento.
- Uso de arquivos convencionais pode ser indicado:
 - ❑ BD simples, bem definidas e sem previsão de mudanças;
 - ❑ Requisitos severos de eficiência computacional;
 - ❑ Acesso de múltiplos usuários não é necessário.

Um pouco sobre a evolução dos SGBD...

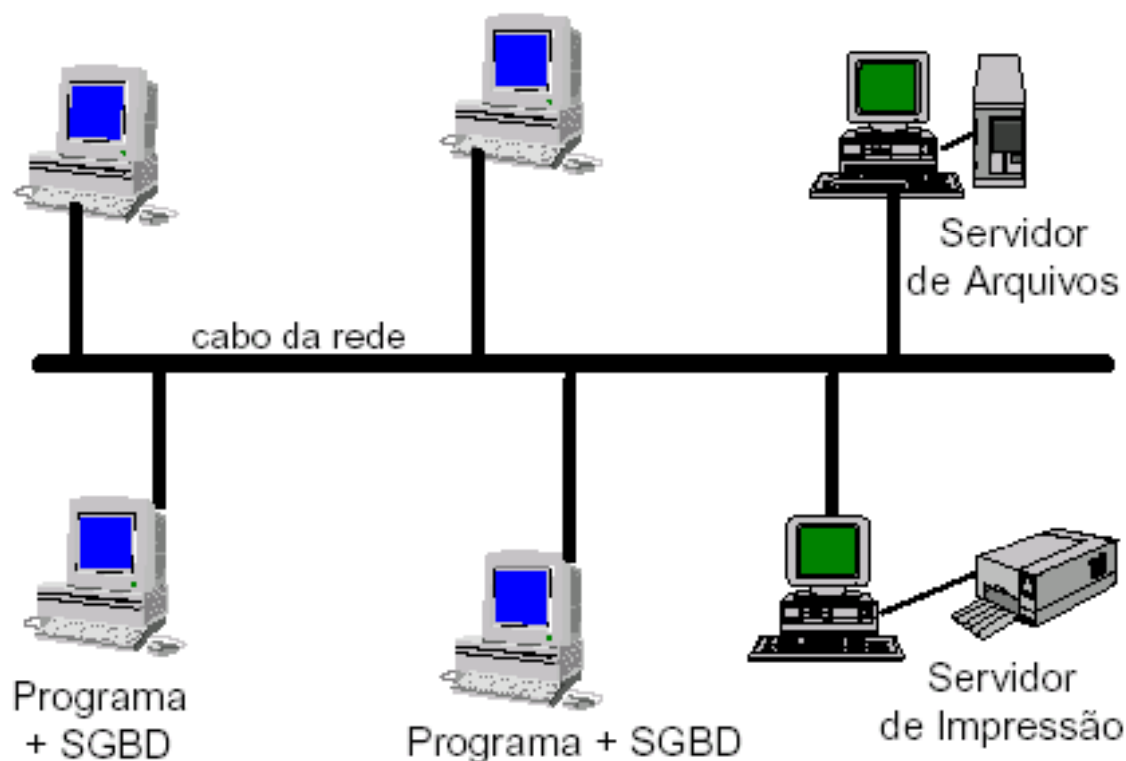
Bancos de Dados Centralizados

(Década de 70)



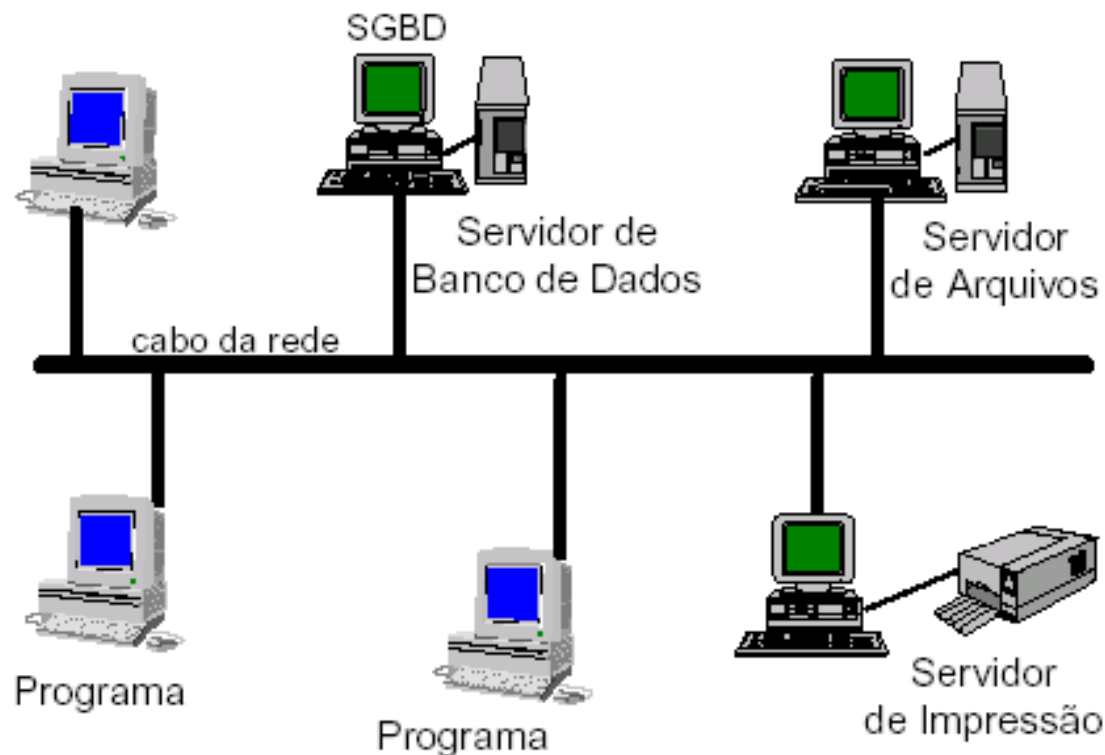
Um pouco sobre a evolução dos SGBD...

Arquitetura Cliente/Servidor com Servidor de Arquivos

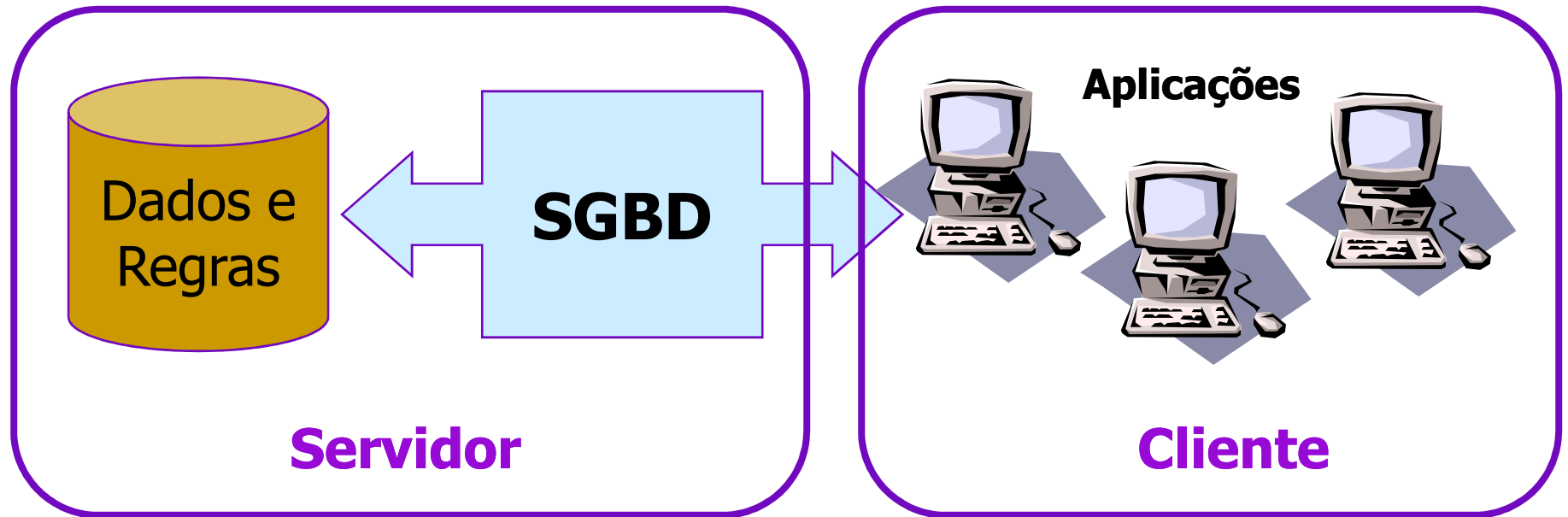


Um pouco sobre a evolução dos SGBD...

Arquitetura Cliente/Servidor com Servidor de Bancos de Dados



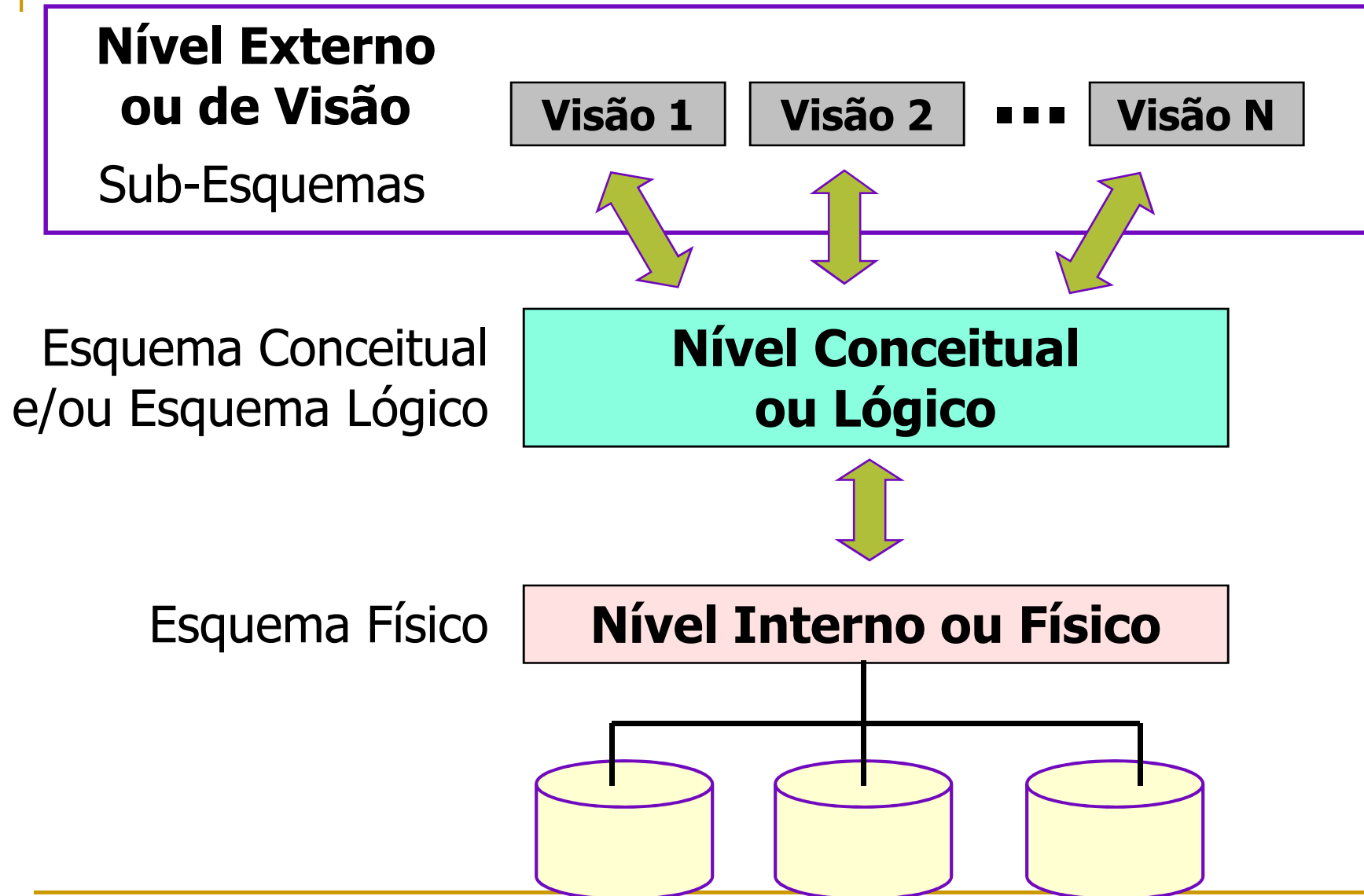
Arquitetura Cliente/Servidor:



Arquitetura de 3 Níveis / Independência de Dados

- Relembrando características relevantes de um BD:
 - ❑ Separação entre programas e dados;
 - ❑ Suporte às múltiplas visões;
 - ❑ Uso de catálogo/dicionário de dados (metadados) para armazenar a descrição do BD (esquema);
- Arquitetura de três níveis auxilia na realização e na visualização dessas características:
 - ❑ Separar o usuário da aplicação do BD físico.

Arquitetura de Três Níveis:



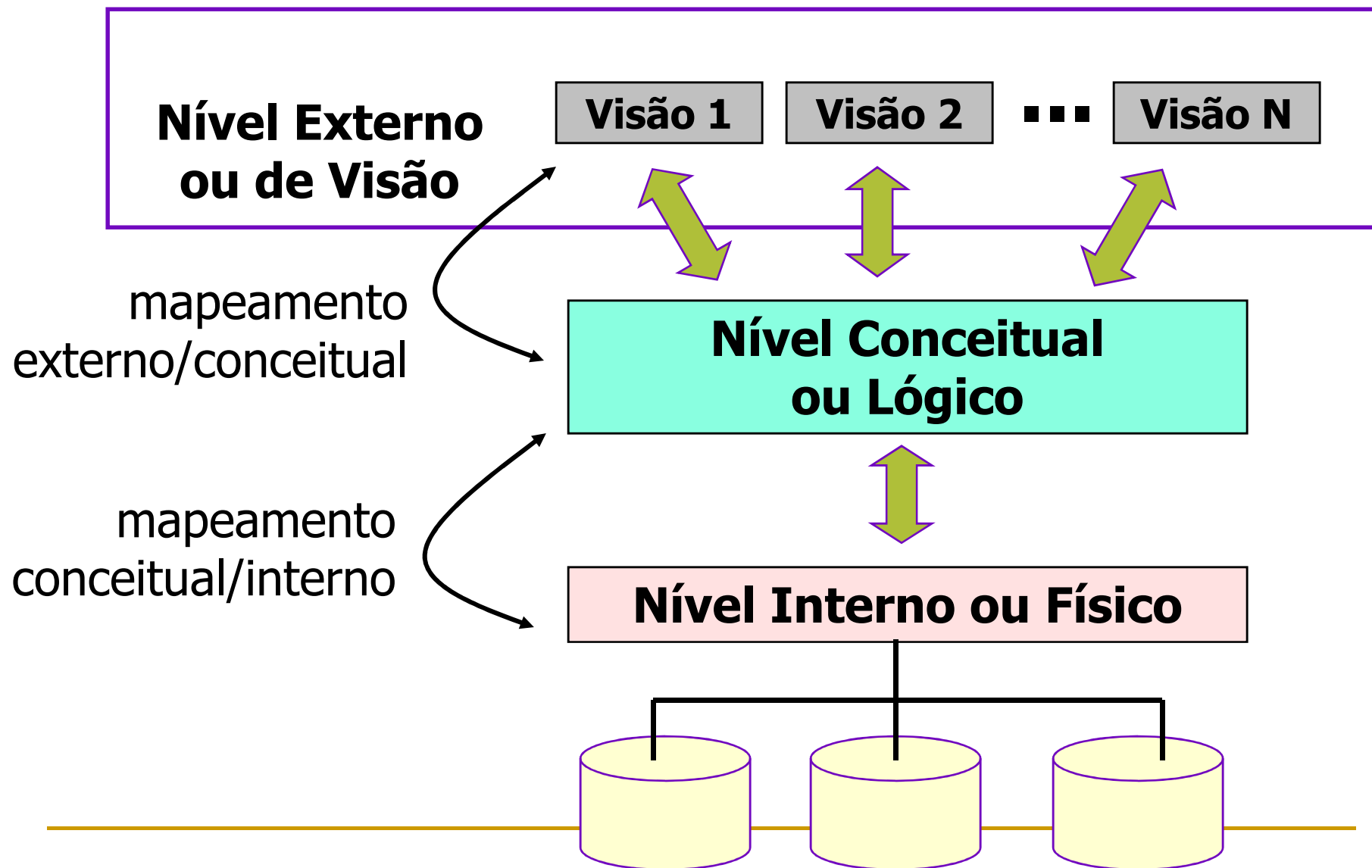
Arquitetura de Três Níveis

- **Objetivos:**
 - ❑ separar as aplicações dos usuários do BD físico;
 - ❑ prover uma visão abstrata dos dados.
- **Três níveis de abstração para organizar os dados:**
 - ❑ organização física (esquema interno);
 - ❑ organização lógica global (esquema conceitual);
 - ❑ organização lógica particular (esquema externo, visão).

Arquitetura de Três Níveis

- Esquema interno:
 - Descreve detalhes de armazenamento (físico);
 - Contém definições de estruturas de dados e mecanismos de acesso.
- Esquema conceitual:
 - Descreve a estrutura do BD (entidades, tipos de dados, etc.);
 - definição do conteúdo da informação;
 - utiliza o conceito de modelo de dados;
 - independe de estruturas de dados e mecanismos de acesso;
- Esquema externo:
 - Descreve somente a parte do BD pelo qual um grupo de usuários tem interesse - usuário vê apenas parte dos dados;
 - As *visões* são também chamadas de *sub-esquemas*.

Mapeamentos



Abstração

- Visualização de níveis de esquema em sistemas de banco de dados;
- *Ocultar* detalhes e complexidade nos diferentes níveis;

Abstrair: desconsiderar detalhes da informação disponível (simplificar) para ter uma visão mais geral (mais abstrata).

Observações

- Maior parte dos SGBD não separa os 3 níveis completamente;
- Esquemas são apenas descrições dos dados;
- Dados que existem de fato estão no nível físico;
- Interfaces:
 - permitem a comunicação entre dois níveis subjacentes
 - consistem em mapeamentos ou transformações
 - nível físico \longleftrightarrow nível conceitual
 - nível conceitual \longleftrightarrow nível externo

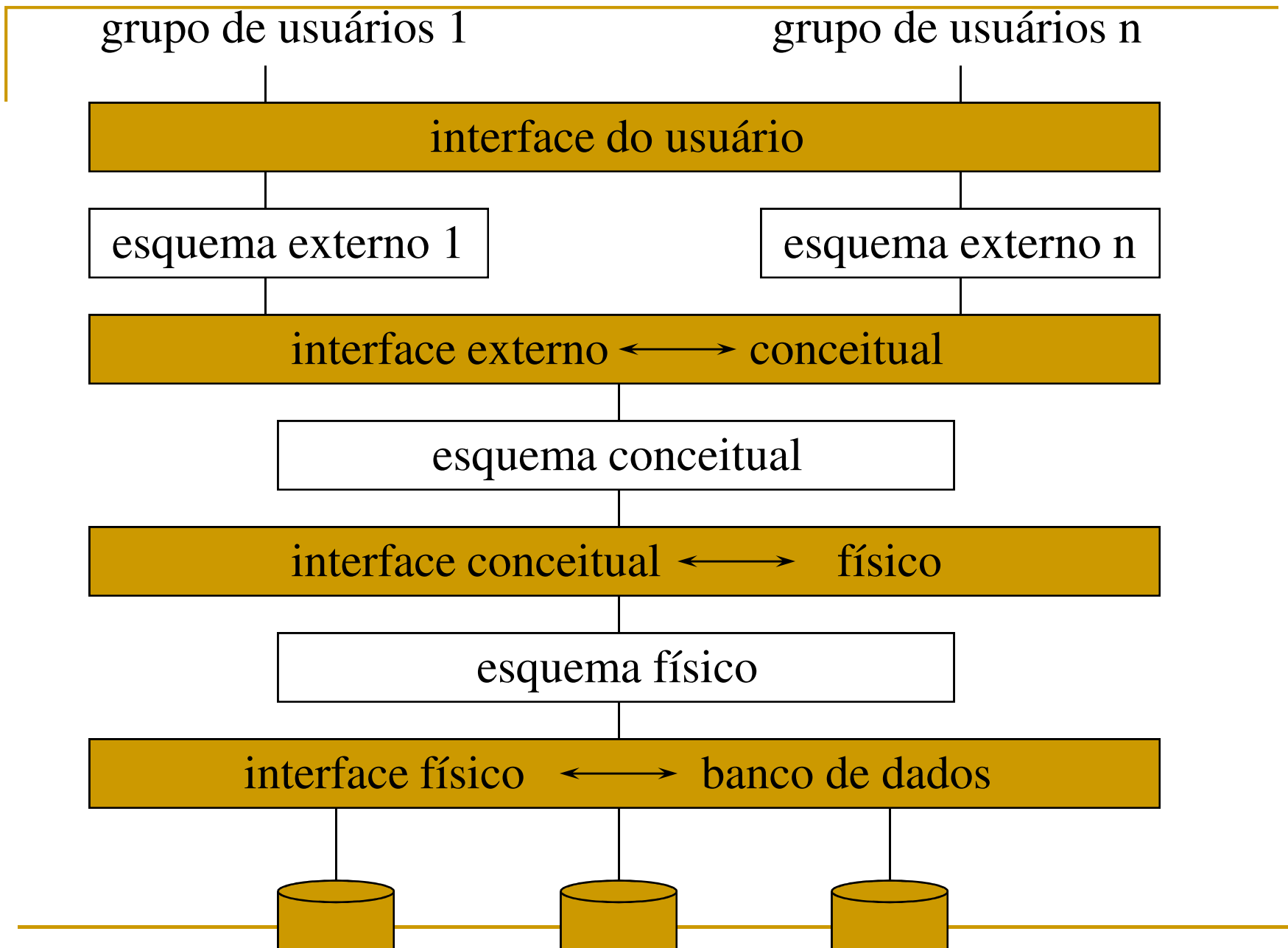
Independência de Dados

- Habilidade de modificar a definição de um esquema em um nível sem afetar a definição do esquema em um nível mais alto;
 - Dois tipos:
 - independência física de dados;
 - independência lógica de dados.
 - Independência física de dados
 - Ao modificar o esquema físico:
 - não modifica os esquemas conceitual e, conseqüentemente, o esquema externo;
 - necessidade: aprimoramento do desempenho.
-

Independência de Dados

- Independência lógica de dados
 - Ao modificar o esquema conceitual:
 - Não modifica os programas aplicativos;
 - Necessidade: alteração da estrutura do BD.

* Alteração do esquema em algum nível implica em alterar os mapeamentos (os programas vinculados ao esquema do nível mais alto não precisam ser alterados).



Modelo de Dados

- BD permitem a abstração dos dados, ocultando detalhes de armazenamento desnecessários para a maioria dos usuários;
- Um *modelo de dados* é um conjunto de conceitos usados para descrever a *estrutura* de um BD;
- *Estrutura*: tipos de dados, relacionamentos e restrições que devem suportar os dados;
- A maioria dos *modelos* inclui uma série de operações básicas para recuperações e atualizações;

Instâncias e Esquemas

■ Instância:

- ❑ coleção de informações armazenadas no BD em um determinado momento;
- ❑ também chamado de extensão do BD;
- ❑ sofre alterações constantemente.

■ Esquema:

- ❑ Descrição do BD (projeto), incluindo as entidades e os relacionamentos entre estas;
 - ❑ também chamado de intenção do BD;
 - ❑ não sofre alterações com frequência.
-

Estado do Banco de Dados

- Os dados armazenados em um BD em um determinado momento:

- Estado vazio:
 - após a criação do BD.
- Estado inicial:
 - após o povoamento (ou carregamento) do BD com os *dados iniciais*.

- Novo estado:
 - após cada operação realizada no BD.
- Estado atual:
 - estado do BD em um determinado momento.

Linguagens associadas oferecidas pelo SGBD

- Para uma verdadeira arquitetura de 3 camadas:
 - Linguagem de definição de visões (*VDL*);
 - Linguagem de definição de dados (*DDL*);
 - Linguagem de definição de armazenamento (*SDL*).
- Entretanto, na maioria dos SGBDs a *DDL* é usada para definir os esquemas conceitual e externo, ou seja, é usada para criar o BD;
- Após compilar o esquema, o BD é instanciado e o usuário precisa de uma linguagem para manipulá-lo: Linguagem de manipulação de dados (*DML*).
- Linguagem SQL: combina *VDL*, *DDL* e *DML*.

Exemplo de *DDL*:

- Criação de uma relação contendo informações pessoais sobre alunos;
- Em SQL:
 - create table
 - alter table
 - drop table
- ...

```
CREATE TABLE aluno ( matrícula NUMBER (10,2),  
                       nome VARCHAR(50),  
                       endereço VARCHAR(50),  
                       data_nascimento DATE )
```

Linguagem de Manipulação de Dados (DML)

- Viabiliza acesso (manipulação) dos dados de maneira compatível com o modelo de dados;
 - Recuperação de dados armazenados
 - Inserção de novos dados
 - Remoção de dados
 - Modificação de dados
- Exemplos em SQL:
 - *Insert* ;
 - *Select* ;
 - *Delete* ;
 - *Update* ; ...

Classificação dos SGBD:

- Existem vários critérios que podem ser usados para classificar um SGBD, dentre os quais iremos abordar:
 - Modelo;
 - Número de usuários;
 - Número de nós;

Classificação de acordo com o modelo

- ❑ Modelo de rede (década de 60);
- ❑ Modelo hierárquico (década de 60);
- ❑ Modelo relacional (a partir da década de 70)*:
 - dados e relacionamentos: coleções de tabelas;
 - cada tabela: várias colunas e nome único;
 - relaciona os registros por meio de valores;
 - desenvolvimento de fundamentos matemáticos para sua definição (Cálculo e Álgebra Relacional).

* Codd, E.F., “A Relational Model for Large Shared Data Banks”,
Communications of the ACM, 13 (6), 377-397, 1970.

BD Relacional:

c l i e n t e	nome	rua	cidade	número
	Adriana	Rua Juca Filho	Maringá	900
	Ricardo	Av. Brasil	Rio de Janeiro	556
	Ricardo	Av. Brasil	Rio de Janeiro	647
	Samanta	Av. Tiradentes	São Paulo	801
	Samanta	Av. Tiradentes	São Paulo	647

c o n t a	número	saldo
	900	55
	556	10000
	647	105366
	801	10533

Outros modelos de dados incluem:

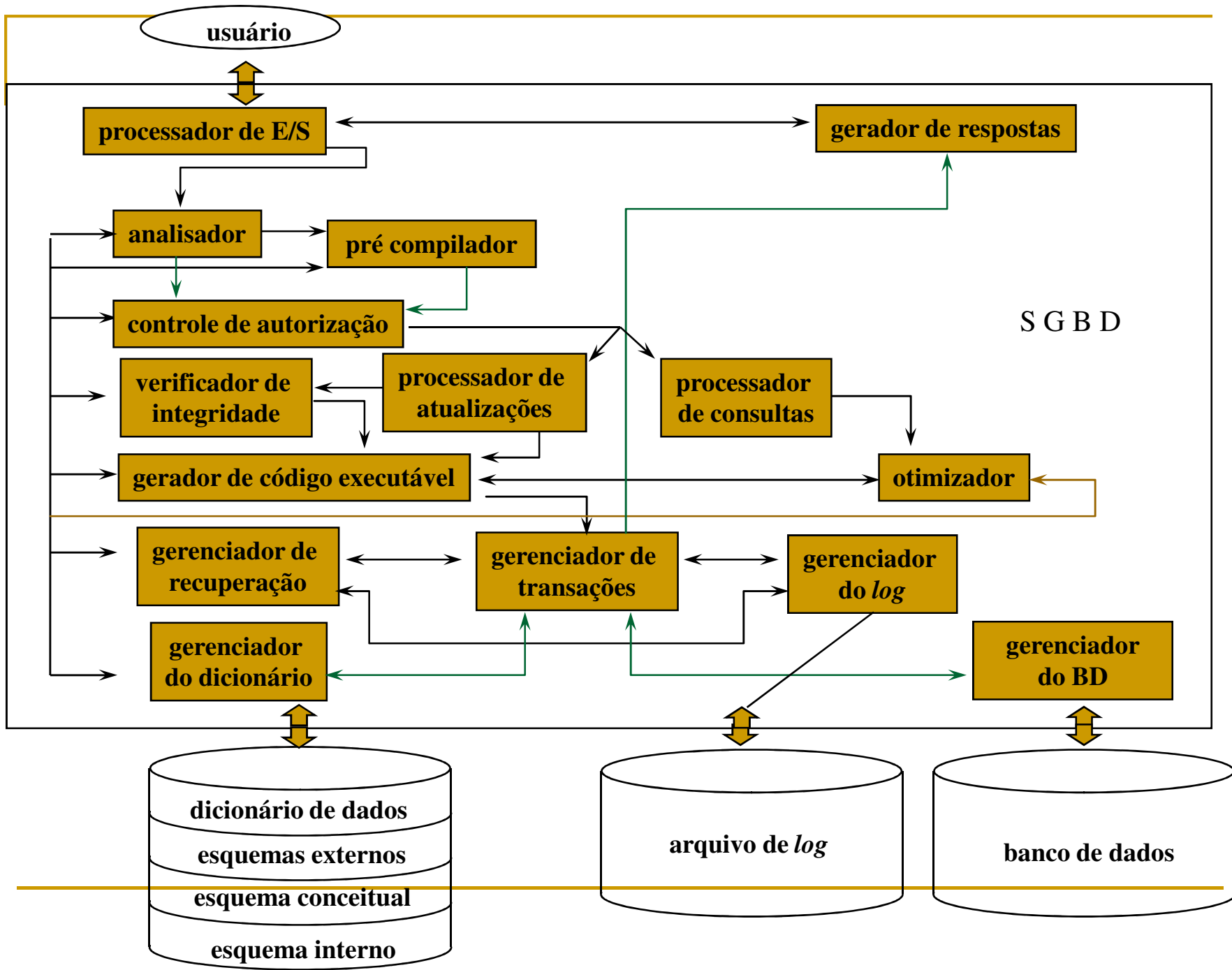
- ❑ Modelo orientado a objetos (década de 90):
 - dados e relacionamentos são coleções de objetos;
 - objeto: estrutura (propriedades) + operações (métodos).
- ❑ Modelo objeto-relacional (década de 90):
 - fundamentado no modelo relacional;
 - estendido com características do modelo orientado a objetos.

Classificação dos SGBD...

- De acordo com o número de usuários:
 - monousuário: um único usuário por vez;
 - multiusuário: vários usuários ao mesmo tempo.
- De acordo com o número de nós:
 - centralizado: dados e SGBD localizados em um único nó;
 - distribuído: dados e SGBD localizados em vários nós, conectados através de redes de comunicação.

Arquitetura de um SGBD

- componentes (processos);
- funcionalidades dos componentes;
- interação existente entre tais componentes.
- Objetivos:
 - enfatizar quais funcionalidades devem ser oferecidas internamente por um SGBD;
 - como estas funções cooperam logicamente ou dependem uma das outras.



Desenvolvimento de SBD

■ Projeto conceitual:

□ esquema conceitual para a base de dados:

- níveis lógico e externo;

- baseado nos requisitos de dados;

- objetivos:

 - Estrutura da base de dados;

 - Semântica;

 - Relacionamentos;

 - Restrições;

- descrição clara, não ambígua e padronizada.

→ modelo conceitual que independe do SGBD (e.g., MER).

Desenvolvimento de SBD...

■ Projeto lógico

- ❑ esquema lógico: níveis lógico e externo.
- ❑ mapeamento do modelo conceitual para o modelo do SGBD (e.g., Modelo Relacional).
- ❑ **Passo 1** – mapeamento independente de um SGBD específico, mas dependente do “paradigma” (e.g., relacional).
- ❑ **Passo 2** – ajustes de acordo com as características e restrições do modelo implementado por um SGBD específico.

Desenvolvimento de SBD...

■ Projeto físico

□ esquema físico:

- nível interno;

□ estruturas físicas de armazenamento:

- organização de registros físicos;
- índices.

□ critérios:

- tempo de resposta;
- espaço utilizado;
- número de transações.

Modelagem de dados - Três Reinos:

Real



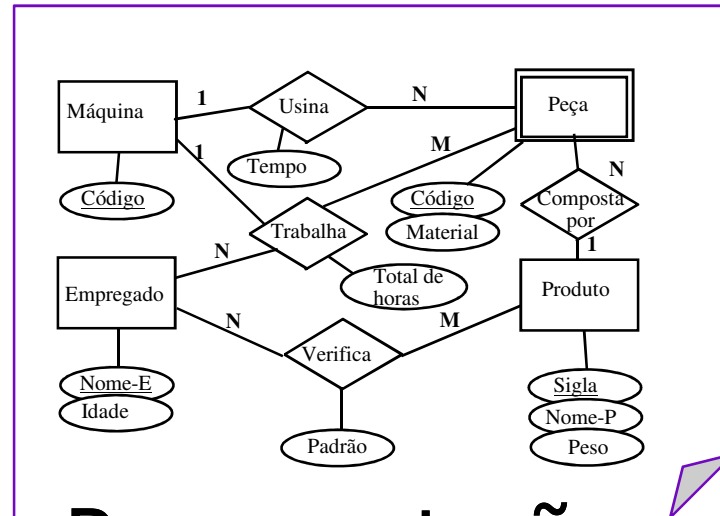
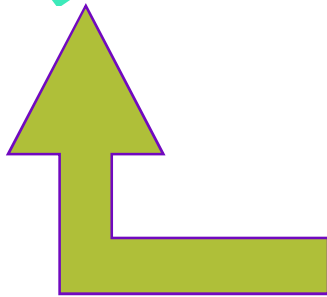
Percepção



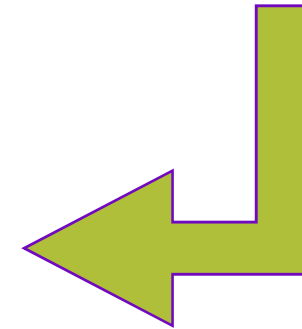
Imaginário



Implementação



Modelagem



Representação

Modelagem de Dados

- Representação da informação do mundo real:
 - esconde detalhes de implementação;
 - estrutura definida.
- **Modelo de Dados** \Rightarrow conjunto de conceitos que permitem definir a estrutura da informação.



mecanismo de **Abstração de Dados**

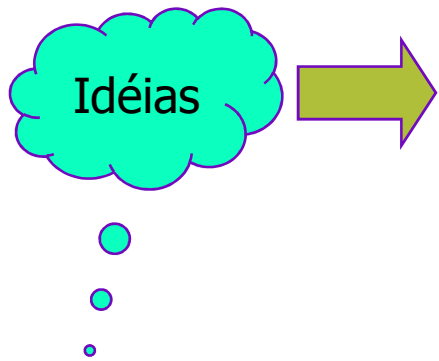
Modelagem de Dados - Motivação

- Em geral: projetistas se apóiam pouco em metodologias sistemáticas para conduzir o projeto da base de dados;
- Conseqüências:
 - tempo e recursos são subestimados;
 - resultado não atende às necessidades das aplicações;
 - documentação é limitada;
 - manutenção custosa;
 - falta compreensão sobre a informação a ser armazenada e manipulada;

Modelos de Dados

- Modelo de dados - definição abstrata, autônoma e lógica dos **objetos**, **operadores** e outros elementos que, juntos, constituem a máquina abstrata com a qual os usuários interagem”. (Date)
 - ❑ objetos - permitem modelar a **estrutura da informação (dados)**;
 - ❑ operadores - permitem modelar o **comportamento** dos dados.

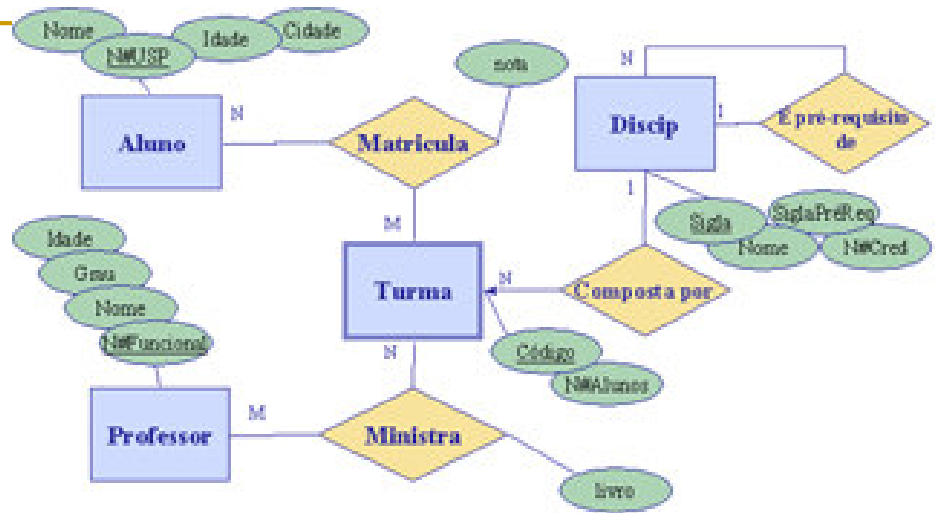
 - Tipos de modelos de dados:
 - ❑ Conceituais:
 - **Modelo Entidade Relacionamento (MER)**;
 - Modelo de Objetos da ODMG (*Object Model*).
 - ❑ de Implementação (Elmasri):
 - Rede, Hierárquico, **Relacional**.
-



Modelo E/R

Modelo Relacional

SGBD Relacional

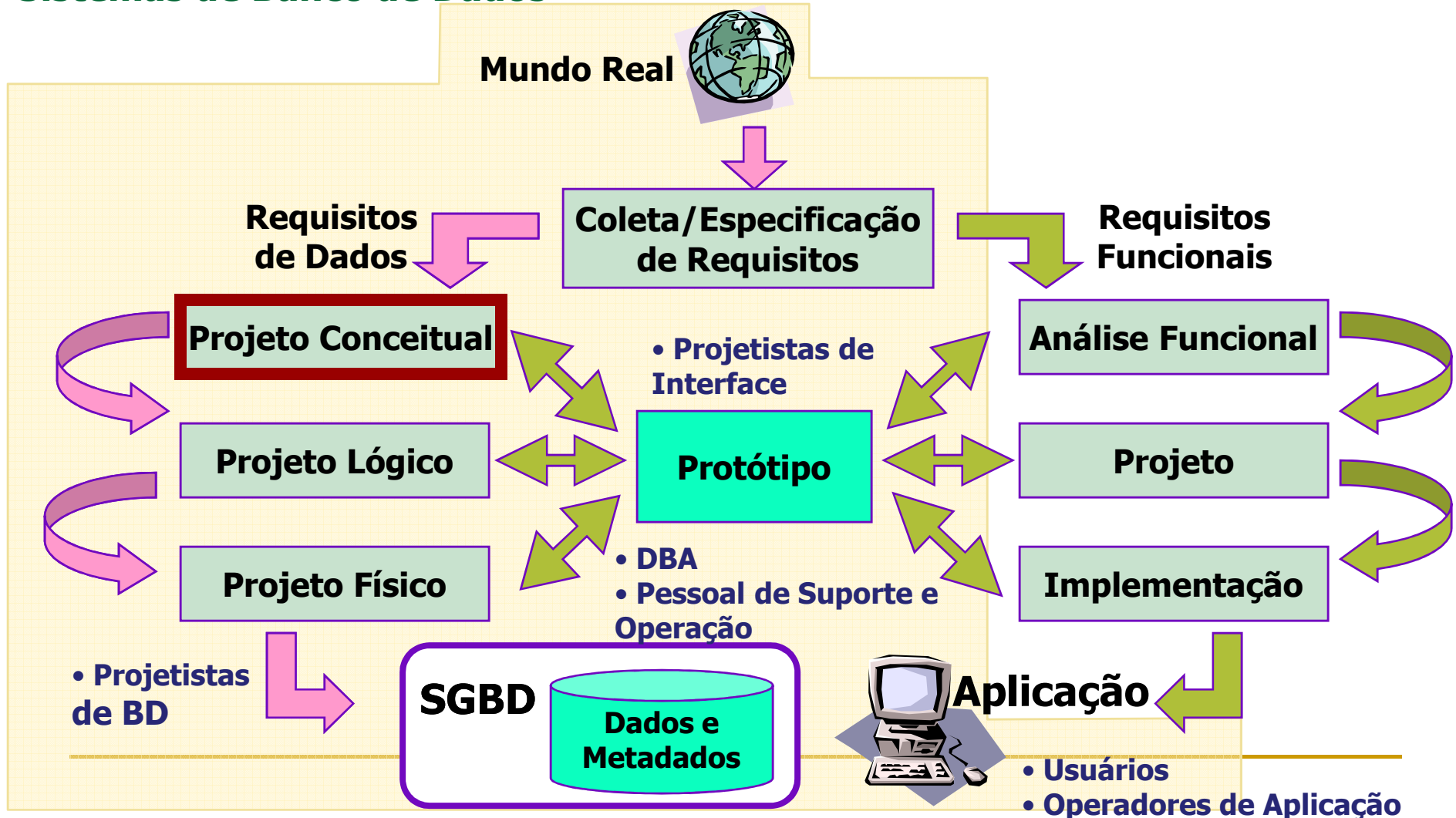


```
Professor = (Nome, N#Funcional, Grau, Idade)
Aluno = (Nome, N#USP, Idade, Cidade)
Turma = (Sigla, Número, Código, N#Alunos)
Discip = (Sigla, Nome, SiglaPréReq, N#Cred.)
Matricula = (CódigoTurma, N#USP, nota)
Ministra = (N#FuncProf, Código, livro)
```

NOME	NUSP	IDADE	CIDADE
Cesar	123	21	Araraquara
Carlos	1234	21	Sao Carlos
Celso	2345	22	Sao Carlos
Cicero	3456	22	Araraquara
Carlitos	4567	21	Ibitinga
Calatana	5678	23	Sao Carlos
Obete	6789	21	Araraquara
Corina	7890	25	Rio Claro
Celina	8901	23	Sao Carlos
Celia	9012	20	Rio Claro
Pedro	10000	19	Sao Paulo
Paulo	10001	19	Curitiba
Joao	54534	23	rub

Projeto conceitual de um SBD

Desenvolvimento de Sistemas de Banco de Dados



Modelagem Conceitual

- Objetivo:

- **descrição do conteúdo** da base de dados;
 - NÃO considera estruturas de armazenamento.

- Enfoque:

- compreensão e descrição da realidade (informação);
- compreensão e seleção das propriedades relevantes da informação;
- compreensão e descrição das restrições sobre os dados;
- diálogo com o usuário.

Modelagem Conceitual

- Entrada: **Requisitos de Dados**

- ao contrário do ideal, a descrição em geral é: informal, dúbia, incompleta, redundante, contraditória, incoerente, longa e de difícil manipulação.

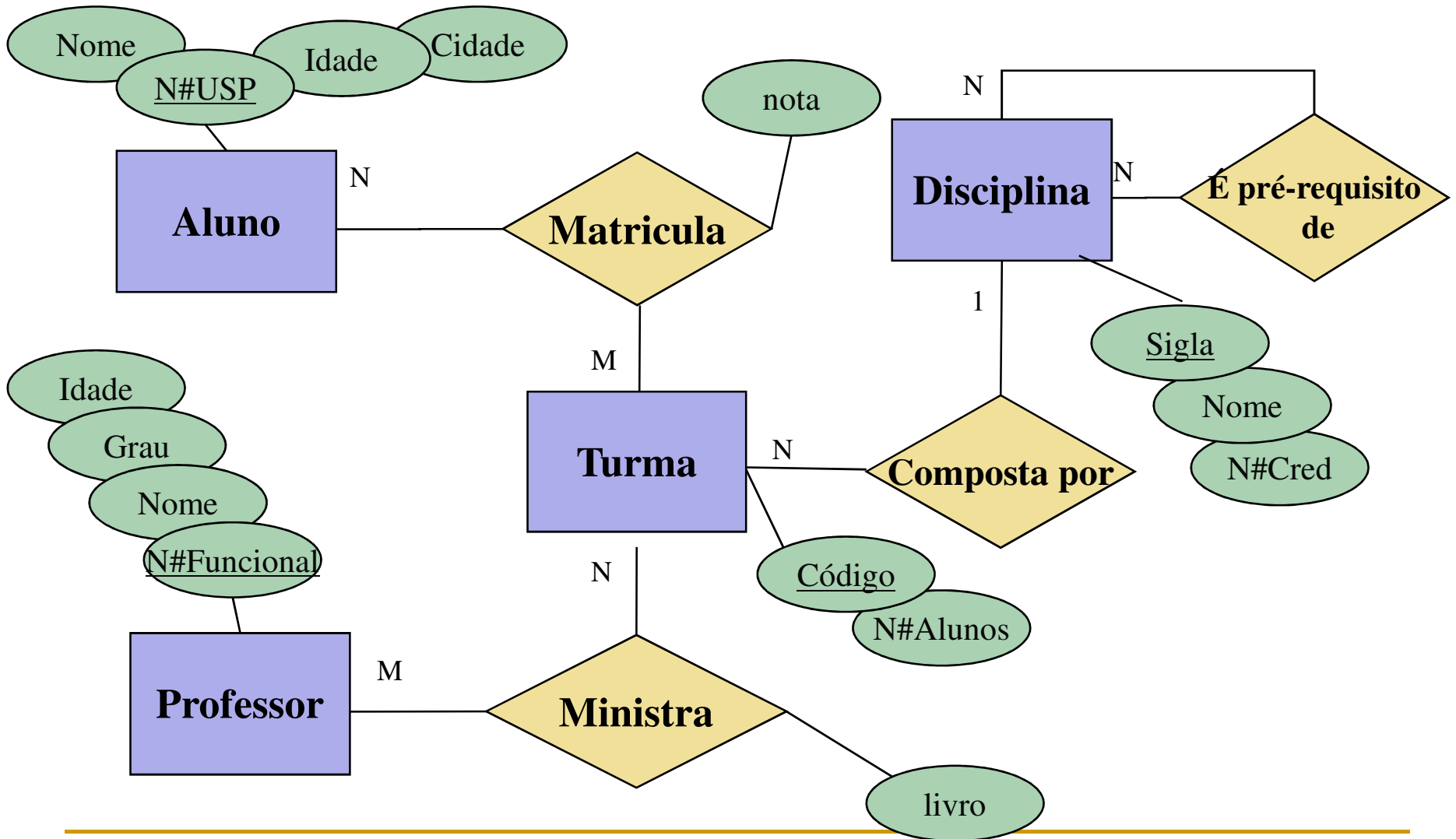
- Processo:

- modelagem – representação conceitual;
- modelo conceitual (Ex: MER);

- Resultado: **Esquema Conceitual**

- descrição sucinta (diagramas e texto);
 - clara, concisa, sem ambigüidades, sem contradições;
 - padronizada.
-

Modelagem Conceitual usando MER:



Modelagem Conceitual

- Vantagens:
 - Facilita a compreensão da informação;
 - Facilita comunicação entre partes envolvidas:
 - Preciso, simples, padrão;
 - Não envolve considerações de implementação;
 - Completo;
 - Coerente: sem ambigüidades e contradições;
 - Representação diagramática - resalta os aspectos mais importantes.
 - Facilidade de manutenção da documentação;
 - Independência de uma representação específica;
 - escolha do SGBD pode ser postergada;
 - esquema conceitual “sobrevive” a alterações de SGBD.

- Próxima aula: modelos conceituais de dados ...
